Access to airports: a case study for the San Francisco Bay Area[1].

First Draft

Eric Pels[23], Peter Nijkamp, Piet Rietveld

Free University Amsterdam, Department of Regional Economics,

Boelelaan 1105, 1081 HV Amsterdam, Netherlands.

Abstract

In this paper (nested) logit models that describe the combined access mode-airport-airline choice are estimated. Nested models with the choice sequence (1) airport-access mode combination and then (2) airline are preferred for Bay Area residents and vistting business travelers. For visiting leisure travelers, a multinomial logit model is prefered.

Keywords: airport choice, access modes, discrete choice models

[1] Preliminary version, comments welcome!

[2] Email:apels@econ.vu.nl. Fax:+31-20-4446004

[3] Also affiliated with the Tinbergen Institute, Keizersgracht 482, 1017 EG Amsterdam.

1 Introduction

Airports in a multiple airport region, where passengers (traveling to a fixed destination) are able to choose between different departure airports, will compete with substitute airports for origin (and also destination) passengers. Passengers take a number of decisions; they have to choose the (departure) airport, airline and airport access mode (given that they already have decided they will fly). These choices can be made simultaneously or sequentially. These choices depend on a number of variables such as airport taxes and airport access times, frequency of service offered by the airline and airfare, and availability and cost of the access mode. Moreover, these choices may be mutually dependent; depending on whether these choices are taken sequentially or simultaneously. The choice of airport access mode has been studied only sparsely in the literature; notable exceptions are Bondzio (1996) and Harvey (1986), who offer empirical studies of the passengers' choice of access mode in Germany and the San Francisco Bay Area respectively. Both authors use multinomial logit models with access time and cost as explanatory variables. Moreover, Bondzio (1996) also estimates nested logit models and finds that business travelers make the choices of access mode and airport sequentially while leisure travelers make the choice simultaneously; the access time to the airport (which is highly mode dependent) is concluded to be higher for business passengers than for leisure travelers.

In most studies on airport choice, the (aggregated) frequency is an explanatory variable, but the airline choice is considered as given. Pels et al. (1998) found that a nested model in which first an airport chosen and then an airline is statistically superior. In this paper we will analyze the choice of access mode in the San Francisco Bay Area in relation to the choice of airport and airline by means of a nested logit model. The results will then be compared to the results (for Germany) of Bondzio (1996) -in order to examine whether there are regional (international) differences in the determinants of access mode choice- and Harvey (1986), addressing the same problem using data for 1980 -in order to investigate whether the determinants of access mode choice have changed over time.

2 The econometric model

In this Section an econometric model for the joint access mode-airport-airline choice will be formulated. First, in Subsection 2.1, a concise review of some of the references already mentioned in the introduction will be offered, while next, based on the discussion of these studies, the econometric model will be formulated in Subsection 2.2.

2.1 Literature review

Air travelers have to make a number of decisions. Access mode, departure airport and airline are but some of the decisions to be made. These decisions are dependent on one another, and should be modeled as such.

Bondzio (1996) estimated nested logit models to explain the joint access mode-departure airport choice for German airports, using access time to the airport, access costs and frequency of service as explanatory variables. For business travelers, a nested model with the choice sequence (1) access mode and then (2) departure airport appeared to be the (statistically) preferred model. For leisure travelers, it was concluded that the nested structure which best replicated the behavior of business travelers did not add much compared to a multinomial logit model, which was therefore the preferred model specification.

Pels et al. (1998) found that a nested logit model with the choice sequence (1) airport, and then (2) airline best explained the joint airport-airline choice for both business and leisure travelers in the San Francisco Bay Area.

In both studies mentioned above, the nested structures were found to be superior to the conventional multinomial logit models. In case of the airport-airline choice, airlines operating from the same airport are closer substitutes than airlines operating from different airports. Hence the introduction of a new airline at a certain airport will affect the airlines already operating from that airport more than airlines operating from alternative airports. Likewise, if the train is the most preferred access mode, airports that can be reached by train are closer substitutes than airports that cannot be reached by train.

Based on the findings of the studies discussed concisely in this subsection, it is expected that a nested logit model would best explain the access mode-airport-airline choice. A corresponding model will be formulated in the next Subsection.

2.2 The econometric model

Suppose a traveler has decided to fly to a particular destination (airport). The traveler then has to choose an airline (*l*), departure airport (*d*) and airport access mode (*a*). There are several alternative model specifications. The most simple one is the multinomial logit model where all combinations (*l,d,a*) are treated as alternatives of which the derived utilities, by assumption, are independent. As a result, if one alternative is added, *all* other alternatives would suffer proportionally; thus if at airport *d* a new airline is introduced, all other alternatives (i.e. also at the alternative airports *d'*) would suffer proportionally, whereas it would be more reasonable to assume alternatives including airport *d* would suffer more. To overcome this "independence of irrelevant alternatives" assumption, a nested multinomial logit model can be specified. Then one recognizes that there are clusters of alternatives of which the derived utilities are correlated. Utilities of alternatives from different classes are not correlated. Then the problem is to identify the different relevant clusters.

Let there be *L* airlines, *D* airports and *A* access modes. The alternatives made up by the airlines operating from the same airport and the airport can be seen as clusters of alternatives: $L(d) \subset L,\ d \in D$ (see Pels et al., 1998). Likewise, the alternatives constituted by the airports and the same access mode can be seen as clusters: $D(a) \subset D,\ a \in A$ (see Bondzio, 1996). The corresponding probability model is:

$$\mathrm{P}(l,d,a) = \mathrm{P}(a)\ \mathrm{P}(d|a)\ \mathrm{P}(l|d,a) \tag{1}$$

$$\mathrm{P}(l|d,a) = \frac{\exp(\frac{V_l}{\mu})}{\sum_{l'} \exp(\frac{V_{l'}}{\mu})}, l \in L(d), d \in D(a), a \in A \tag{2}$$

$$P(d|a) = \frac{\exp(\frac{V_d + \mu \ln \sum_l \exp\left(\frac{V_l}{\mu}\right)}{\theta})}{\sum_{d'} \exp(\frac{V_{d'} + \mu \ln \sum_l \exp\left(\frac{V_l}{\mu}\right)}{\theta})}, d \in D(a), a \in A \quad (3)$$

$$P(a) = \frac{\exp(V_a + \theta \ln \sum_d \exp\left(V_d + \mu \ln \sum_l \exp\left(\frac{V_l}{\mu}\right)\right))}{\sum_{a'} \exp(V_{a'} + \theta \ln \sum_d \exp\left(V_d + \mu \ln \sum_l \exp\left(\frac{V_l}{\mu}\right)\right))}, a \in A \quad (4)$$

where V_l is the systematic utility derived from airline l, V_d is the systematic utility derived from airport d and V_a is the systematic utility derived from access mode a. In this model, a passenger chooses an access mode based on characteristics of the access mode and the maximum expected utility of using the airports available when using the access mode. The passenger chooses an airport based on characteristics of the airport and the maximum expected utility of using the airlines available from the airport. In other words, the passenger first chooses the access mode, then the departure airport and then the airline. Alternatively, the passenger chooses the access mode and departure airport simultaneously and then chooses the airline: $L(d,a) \subset L$, $d,a \in D \times A$ and in the probability model $\theta = 1$.

Let the systematic utility of using an airline l be given by

$$V_l = \alpha_p p_l + \alpha_f \ln(f_l) + \alpha_s \ln(s_l) \quad (5)$$

where p_l is the airfare charged by airline l; $\alpha_p < 0$. f_l is the frequency of service, included in logarithmic form, as it is an indication of the "size" of an airline in a market to a certain destination; $\alpha_f > 0$. s_l is the average number of seats, included in logarithmic form, as it is also an indication of the "size" of an airline[4]. Moreover, aircraft size can be seen as an indicator of the level of comfort; larger aircraft have more amenities. We use the average number of seats as a proxy for aircraft size. To account for decreasing marginal utility of comfort, it is in logarithmic form; $\alpha_s > 0$. The systematic utility of using an airport d is given by

[4] The "size" of an airline in an origin-destination market can be represented by $S_l = f_l s_l$. S_l is best included in logarithmic from in the utility function; see Ben-Akiva and Lerman (1987, chapter 9) for details.

$$V_d = \beta_d + \beta_t d_d \tag{6}$$

where β_d is an airport specific constant, d_d is the road distance to the airport; $\beta_t < 0$. The systematic utility of using an access mode a is

$$V_a = \gamma_a + \gamma_p p_a + \gamma_t t_a + \gamma_c c \tag{7}$$

where γ_a is a mode specific constant, p_a is the cost of the access mode, $\gamma_p < 0$. t_a is the access time to the airport using access mode a; $\gamma_t < 0$ and c stands for personal characteristics (such as group size, pieces of luggage etc.). $\mu < \theta < 1$. When $\mu = \theta = 1$, the model reduces to the multinomial logit model.

3 The 1995 MTC Airline Passenger Survey

The 1995 Metropolitan Transportation Commission Airline Passenger was conducted in August and October 1995 at San Francisco International Airport (SFO), San Jose International Airport (SJC), Oakland International Airport (OAK) and Sonoma County Airport (STS). Some 21,500 passengers departing from these airports were interviewed within 45 minutes to 1 hour before take off; see Table 1 for the distribution of respondents over the airports.

Table 1 Respondents and total enplaned passengers (1995)

Airport	San Francisco	San Jose[1]	Oakland	Sonoma County	Total
Respondents	10,454	7,119	3,497	54	21,124
Passengers	15,013,265	4,267,071	7,750,857	<500,000	

1) A disproportionally large number of interviews was conducted at San Jose at the request of the airport authorities.

In Table 2 the distribution of respondents over the different access modes is given for each of the four airports. The majority of the passengers, both business and leisure, use a car to get to the airport. The percentage of private cars used by visiting passengers is quite high. This can only be the case if some (most) of these passengers are dropped off (at the airport). This information is not available; the information that is available is the number of people that came into the terminal to see the respondent off. Hence one can only expect that more respondents (especially visitors) were dropped off at the airport and *not* accompanied into the terminal. This assumption is reinforced by the fact that a visiting business traveler is more likely to use a rental car than a private car, while for visiting leisure travelers the reverse holds true.

Compared to SJC, OAK and STS, passengers at SFO use more often the access modes that are alternatives to the car (though the car, whether private or rented, is by far the most likely access mode).

Table 2a Shares of different access modes (%), SFO

	residents			visitors		
	business	leisure	total[1]	business	leisure	total[1]
private car	67	61	64 (84)	14	31	24 (68)
rental car	3	2	2 (95)	40	29	34 (96)
private scheduled	6	9	8 (97)	7	7	7 (96)
public transit	1	2	2 (99)	1	2	1 (94)
door 2 door van	11	15	13 (98)	13	15	14 (97)
hotel courtesy	1	3	2 (97)	7	7	7 (98)
taxi	7	5	6 (97)	12	6	8 (97)
limousine	3	2	2 (95)	5	3	4 (97)

1) In brackets the % of respondents using a private car who were *not* accompanied by someone into the terminal to see the respondent off.

Table 2b Shares of different access modes (%), SJC

	residents			visitors		
	business	leisure	total[1]	business	leisure	total[1]
private car	87	88	88 (81)	24	66	43 (66)
rental car	3	1	2 (91)	61	25	46 (94)
private scheduled	1	2	1 (100)	2	2	2 (96)
public transit	0	1	1 (100)	0	1	1 (95)
door 2 door van	2	2	2 (99)	1	1	1 (97)
hotel courtesy	0	0	0 (100)	8	2	5 (98)
taxi	5	5	5 (91)	4	2	3 (97)
limousine	2	1	1 (95)	1	1	1 (81)

1) see footnote at Table 2a

Table 2c Shares of different access modes (%), OAK

	residents			visitors		
	business	leisure	total[1]	business	leisure	total[1]
private car	87	83	84 (79)	32	62	52 (57)
rental car	2	1	1 (96)	50	24	33 (93)
private scheduled	2	3	3 (98)	2	3	3 (92)
public transit	3	7	6 (99)	4	4	4 (90)
door 2 door van	3	4	4 (94)	5	3	3 (100)
hotel courtesy	0	0	0 (100)	2	1	2 (96)
taxi	2	1	1 (88)	4	2	3 (91)
limousine	1	0	1 (100)	0	0	0 (100)

1) see footnote at Table 2a

Table 2d Shares of different access modes (%), STS

	residents			visitors		
	business	leisure	total[1]	business	leisure	total[1]
private car	83	100	88 (89)	46	86	60 (50)
rental car	8	0	8 (50)	36	14	28 (100)
taxi	4	0	3 (100)	9	0	6 (100)
limousine	4	0	3 (100)	9	0	6 (0)

1) see comment made under Table 2a

All airports can be reached using public transportation, but whether public transportation is a likely access mode depends on the city of origin and the airport used. SFO, SJC and OAK can be reached by rail and bus from some cities, and only by bus from other cities. In the analysis, rail and bus are joined in the access mode public transportation; we assume public transportation is available to each passenger. In map 1 the airports, road system and points of origin for the respondents are given.

map 1 about here

4 Estimation Results

To be able to estimate the model, data on travel times and costs for the different access modes are necessary. Using a road map of the San Francisco Bay Area[5], access times using a private car could calculated using the latitude and longitude of the point of origin and the airports in the system. Access times for the other modes were estimated as: access time for the private car + 15 minutes for taxi, door to door van and rental car and twice the access time using a private car for public transportation.

[5] Downloadable from www.bts.gov

For hotel courtesy the same access time as for the private car was used. As the information on whether a respondent using a private car was dropped off or used a (longer term) parking lot was incomplete, the cost of using a private car was fixed at the cost of a 24 hour parking period. It is noted that for some respondents this may be too high while it is too low for others. Based on the price information found on different websites of car rental companies, the cost of using a rental car was set at $50. For the modes taxi, door to door van and public transportation average costs could be found on the web for some city-airport pairs. Based on these data, average access costs per mile could be calculated. These were $2.50 per mile for a taxi, $1.10 per mile for a door to door van and $1 base + $0.05 per mile for public transportation. Hotel courtesy was assumed free of charge.

The model as specified in equations (1)-(7) was estimated using FIML. The full nested structure (with three levels) did not deliver viable results. Moreover, the statistically preferred model for the joint access mode-airport choice (using the aggregated frequency and number of seats as explanatory variables for the airport choice) was a multinomial logit model. Hence the model was restated such that the passenger first chooses an access mode-airport combination and then an airline.

Estimation results for Bay Area residents are presented in Table 3. The available access modes are: private car, rental car, door to door van (including private scheduled)[6], public transportation and taxi (including limousine). In Table 3 rental car is a dummy variable which takes the value 1 if the respondent has chosen a rental car and has used it for other purposes besides driving to the airport. Home is a dummy variable which takes on the value 1 if the respondent's origin was his/her home. As this variable has the same values accross all mode-airport combinations, the parameter for the mode private car was normalized to 0. To avoid multi-collinearity (with the rental car dummy) the parameter for the mode rental car was also fixed at 0. The parameter μ describing the heterogeneity between airlines is made airport specific; $0 < \mu < 1$. Airlines are closer substitutes if μ is closer to 0.

[6] Private scheduled and door to door van are treated as the same access mode, although preferable they should be treated as different access modes. However, for technical reasons and because the average

Table 3 Estimation results, Bay Area residents, August 1995

		Business travelers		Leisure travelers	
Parameter		Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
ln(frequency)		0.85131	0.83694E-01	0.638582	0.67999E-01
ln(seats)		1.4555	0.36331	1.6626	0.28511
constant	SFO	reference		reference	
	SJC	1.0215	0.18486	1.0274	0.16321
	OAK	1.4017	0.11968	1.9070	0.91340E-01
access cost		-0.38181E-01	0.53941E-02	-0.23268E-01	0.44933E-02
access time		-0.88490E-02	0.20696E-02	-0.14344E-01	0.16538E-02
rental car		4.6546	1.1512	4.2217	1.2208
home	car	reference		reference	
	d2d[1]	1.6569	0.32151	1.6352	0.21016
	pt[2]	-0.67739	0.31746	-0.33338	0.20788
	taxi	2.1498	0.44677	1.2405	0.42649
μ_{SFO}		0.37446	0.87101E-01	0.47049	0.75418E-01
μ_{SJC}		0.19857	0.79549E-01	0.36100	0.70909E-01
μ_{OAK}		0.19797	0.79952E-01	0.32423	0.69882E-01
Log-likelihood		-1864.406		-3145.597	
ρ^2		0.46		0.43	
observations		794		1278	

1) door to door van
2) public transportation

From Table 3 it appears that business travelers are more sensitive to frequency, but are less sensitive to access time and access cost than leisure travelers. The latter finding seems to contradict the common finding in the literature that business passengers are more sensitive to access times than leisure travelers. In the "business model" the alternatives within the clusters (clusters constituted of airlines and the same airport-access mode combination) are closer substitutes than in the leisure model. Moreover, the alternatives within the clusters not including SFO seem to be closer substitutes than the alternatives within the clusters including SFO. When a passenger will use a rental car also for other reasons than going to the airport, the rental car is more likely to be chosen as the access mode. When leaving from home, business travelers are more likely to choose a taxi then leisure travelers would. Public transportation is a less likely access mode when leaving home. Given the access times (and the maximum expected utilities of the airlines operating from the airports), passengers seem to prefer SJC and OAK over SFO.

---

costs found on the web were in a number of cases given for private scheduled and door to door van

Estimations for Bay Area visitors are presented in Table 4. The available access modes are: rental car, door to door van, public transportation, taxi and hotel courtesy. For the leisure travelers, the nested structure was rejected: the μ's were larger than 1 and therefore theoretically not valid. Hence for the visiting leisure travelers a multinomial logit model is preferred. Models including both the access time and access cost led to theoretically invalid results: the sign for the access cost took on the wrong value. Hence the estimation results for models with only the access time parameter are presented.

Table 4 Estimation results, Bay Area visitors, August 1995

		Business travelers		Leisure travelers	
Parameter		Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
ln(frequency)		0.83488	0.93899E-01	1.2694	0.98588E-01
ln(seats)		1.2276	0.46122	1.9627	0.40400
constant	SFO	reference		reference	
	SJC	0.77813		1.4608	0.18818
	OAK	0.73498		0.88862	0.13066
access time		-0.16871E-01	0.19857E-02	-0.46228E-01	0.27616E-02
hotel	hotel[1]	reference		reference	
	rent[2]	0.86428	0.15405	1.3923	0.18354
	d2d[3]	-0.61566	0.27081	1.3169	0.26878
	pt[3]	0.35715	0.44644	0.65484	0.32571
	taxi	1.6708	0.18228	0.50547	0.14175
μ_{SFO}		0.29157	0.81419E-01	-	
μ_{SJC}		0.13424	0.70440E-01	-	
μ_{OAK}		0.13363	0.71660E-01	-	
Log-likelihood		-2049.64		-1579.871	
ρ^2		0.42		0.44	
observations		813		656	

1) hotel courtesy
2) rental car
3) door to door van
4) public transportation

The parameters for the frequency and seats in the business model do not differ that much from the business model for the residents: the estimates fall within each other's 95% confidence interval. Like in the model for the residents, the alternatives within the clusters not including SFO seem to be closer substitutes than the alternatives

together, they were treated as the same.

within the clusters including SFO. For the leisure travelers, a multinomial logit model is preferred, in which there are no clusters and no perfect substitutes. Because of the different model structures, the parameters are difficult to compare.

5 Conclusion

In this paper discrete choice models describing the access mode-airport-airline choice were estimated. In a simplified model describing the access mode-aiport choice the nested structure first airport, then airline, which Bondzio (1996) found to be the preferred model for business passengers in Germany, was rejected in favour of a multinomial logit model. Hence the model describing the access mode-airport-airline choice has two levels. First the access mode and airport are chosen simultaneously, based on access mode and airport characteristics and the maximum expected utility from the airlines available from each access mode-airport combination. Next, the airline is chosen. This structure was statistically preferable to the multinomial logit model for both resident business and leisure travelers. For resident passengers (both business and leisure), access times and access costs were significant in the access mode-airport choice. For visiting passengers (both leisure and business) on the other hand, models with the access time were preferred.

An interesting finding is that the alternatives (airlines) available from the clusters including Oakland International Airport or San Jose International Airport are closer substitutes than the alternatives available from San Francisco International Airport. This may be due to the fact that, in general, there are more airlines available to a given destination than from the other three airports.

The following research agenda follows from this paper. First and foremost, more research has to be done to be able to derive more reliable access times and costs. Second, airfares should be included in the analysis.

References

Ben-Akiva, M. and S.R. Lerman (1987), *Discrete Choice Analysis: Theory and Application to Travel Demand*, Cambridge, Mass., MIT Press.

Bondzio, L. (1996), *Modelle fuer den Zugang von Passagieren zu Flughaefen* (Models for the Passengers' Access to Airports), Ph.D. thesis, Ruhr-University Bochum (in German).

Harvey, G. (1987), Airport Choice in a Multiple Airport Region, *Transportation Research*, **21A**(6), 439-449.

Ndoh, N.N., D.E. Pitfield and R.E. Caves (1990), Air Transportation Passenger Route Choice: A Nested Multinomial Logit Analysis, *Spatial Choices and Processes*}, eds Fischer, M.M., P. Nijkamp and Y.Y. Papageorgiou, 349-365, Elsevier Science Publishers, North-Holland.

Pels, E., P. Nijkamp and P. Rietveld (1997), Substitution and Complementarity in Aviation: Airports vs Airlines, *Transportation Research*, **33E**, 275-286.

Pels, E., P. Nijkamp and P. Rietveld (1998), *Airport choice in a multiple airport region: a case study for the San Francisco Bay Area*, Free University Amsterdam, mimeo.

Map 1 The Bay Area Road System, Locations of Airports and Respondents.